Analytic properties of aggregation functions for decision making

Endre Pap

Department of Mathematics and Informatics, University of Novi Sad
Trg D. Obradovića 4, 21 000 Novi Sad, Serbia
e-mail: pape@eunet.yu

Abstract: Aggregation functions are closely related to decision making. There is given a short overview of some basic analytical properties of aggregation functions as continuity and measures. Special attention is taken on the integral representations of aggregation functions.

Keywords: Aggregation function, decision making, continuity, Lipschitz property, Choquet integral, Benvenuti integral.

1 Introduction

The aggregation operators (we shall call them, for real valued case, aggregation functions) form a fundamental part of multi-criteria decision making, engineering design, expert systems, pattern recognition, neural networks, fuzzy controllers, genetic algorithms. In many systems (specially intelligent) the aggregation of incoming data plays the main role. Aggregation functions play important role in many different approaches to decision making [5, 7, 9]. The choice of the aggregation function depends on the actual application. To obtain a sensible and satisfactory aggregation, any aggregation function should not be used. To choose satisfactory aggregation functions, we can adopt an axiomatic approach and impose that these functions fulfill some selected properties. These properties can be dictated by the nature of the values to be aggregated, e.g., in some multi-criteria evaluation methods, the aim is to assess a global absolute score to an alternative given a set of partial scores with respect to different criteria. It would be unnatural to give as global score a value which is lower than the lowest partial score, or greater than the highest score, so that only *internal* aggregation functions are allowed. If preference degrees coming from transitive (in some sense) relations are combined, it is natural to require that the result of combination remains transitive. Another example is related to the aggregation of opinions in voting procedures. Since usually the voters are anonymous, the aggregation function have to be symmetric.

We investigate some properties of aggregation functions, restricting with more details on analytical properties (continuity and measures). We present different type of continuity properties. Decision making needs more general mathematical models, which involve also non-additive measures. Previously used additive probability measures could not model some situations as e.g. the Ellsberg Paradox, see [7].

2 Algebraic properties of aggregation functions

First we present some basic mathematical properties of the aggregation functions based on the book under preparation [8]. Throughout we denote by $\mathbb{I}$ any nonempty real interval, bounded or not. The integer n represents the number of values to be aggregated and $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{I}^n$.

Definition 1 *An n-ary aggregation function in $\mathbb{I}^n$ is a function $\mathsf{A}^{(n)} : \mathbb{I}^n \to \mathbb{I}$ that is nondecreasing, i.e., for any $\mathbf{x}, \mathbf{x}' \in \mathbb{I}^n$,*

$$x_i \leqslant x_i' \ \forall i \in \{1, \ldots, n\} \quad \Rightarrow \quad \mathsf{A}(\mathbf{x}) \leqslant \mathsf{A}(\mathbf{x}'),$$

and $\inf_{\mathbf{x}\in\mathbb{I}^n} \mathsf{A}^{(n)}(\mathbf{x}) = \inf \mathbb{I}$ *and* $\sup_{\mathbf{x}\in\mathbb{I}^n} \mathsf{A}^{(n)}(\mathbf{x}) = \sup \mathbb{I}$.

Usually in the practice we consider the case $\mathbb{I} = [0, 1]$, and then the last two equalities in the previous definition reduces on $\mathsf{A}^{(n)}(0, \ldots, 0) = 0$ and $\mathsf{A}^{(n)}(1, \ldots, 1) = 1$. For example, the *arithmetic mean* as an aggregation function is defined by

$$\mathsf{AM}^{(n)}(x_1, \ldots, x_n) = \frac{1}{n} \sum_{i=1}^{n} x_i.$$

When no confusion can arise, the aggregation function will be written A instead of $\mathsf{A}^{(n)}$. We use the convention $\mathsf{A}^{(1)}(x) = x \quad (x \in \mathbb{I})$. Note that, in that in general for different n and m the functions $\mathsf{A}^{(n)}$ and $\mathsf{A}^{(m)}$ need to be related. T

For the illustration and the next use we now give some well-known aggregation functions. For any weight vector $\omega = (\omega_1, \ldots, \omega_n) \in [0, 1]^n$ such that $\sum_{i=1}^n \omega_i = 1$, the *weighted arithmetic mean* WAM_ω and the *ordered weighted averaging* function OWA_ω associated to ω, are respectively defined by $\mathsf{WAM}_\omega(\mathbf{x}) = \sum_{i=1}^n \omega_i \, x_i$, $\quad \mathsf{OWA}_\omega(\mathbf{x}) = \sum_{i=1}^n \omega_i \, x_{(i)}$. For any $k \in \{1, \ldots, n\}$, the *coordinate projection* P_k and the *order statistic* function OS_k associated to the kth argument, are respectively defined by $\mathsf{P}_k(x) = x_k$, $\quad \mathsf{OS}_k(x) = x_{(k)}$. The projection of the first and the last coordinates are defined as $\mathsf{P}_F(x) = \mathsf{P}_1(\mathbf{x}) = x_1, \mathsf{P}_L(x) = \mathsf{P}_n(\mathbf{x}) = x_n$. Similarly, the extreme order statistics $x_{(1)}$ and $x_{(n)}$ are respectively the minimum and maximum $\mathsf{Min}(\mathbf{x}) = \min(x_1, \ldots, x_n)$, $\mathsf{Max}(\mathbf{x}) = \max(x_1, \ldots, x_n)$. Also, the *median* of an odd number of values $(x_1, \ldots, x_{2k-1})$ is simply defined by $\mathsf{Med}(x_1, \ldots, x_{2k-1}) = x_{(k)}$. The sum and product are respectively defined by $\Sigma(\mathbf{x}) = \sum_{i=1}^n x_i$, $\quad \Pi(\mathbf{x}) = \prod_{i=1}^n x_i$.

Notations: For any integer $k \geqslant 1$ and any $x \in \mathbb{I}$, we set $k \cdot x := x, \ldots, x$ (k times). For any vectors $\mathbf{x}, \mathbf{x}' \in \mathbb{I}^n$, we denote by $\mathbf{x}\mathbf{x}'$ the n-dimensional vector $(x_1x_1', \ldots, x_nx_n')$ obtained by calculated the product componentwise. The vectors $\mathbf{x} + \mathbf{x}'$, $\mathbf{x} \wedge \mathbf{x}'$, and $\mathbf{x} \vee \mathbf{x}'$ are defined similarly.For any $\mathbf{x} \in \mathbb{I}^n$ and any function $\varphi : \mathbb{I}^n \to \mathbb{R}^n$, we denote by $\varphi(\mathbf{x})$ the n-dimensional vector $\big(\varphi(x_1), \ldots, \varphi(x_n)\big)$. For any finite or denumerable set K, we let Π_K denote the set of all permutations on K.Given a vector $(x_1, \ldots, x_n)$ and a permutation $\sigma \in \Pi_{[n]}$, the notation $[x_1, \ldots, x_n]_\sigma$ means $x_{\sigma(1)}, \ldots, x_{\sigma(n)}$, that is, the permutation σ of the indices.

The first property we consider is *symmetry*, also called commutativity, neutrality, or anonymity. The commutativity of binary operation $*$ is usually given in the form $x * y = y * x$, can be easily generalized for n-ary aggregation functions, with $n > 2$, as follows.

Definition 2 $\mathsf{A} : \mathbb{I}^n \to \mathbb{I}$ *is a symmetric function if*

$$\mathsf{A}(\mathbf{x}) = \mathsf{A}([\mathbf{x}]_\sigma) \qquad (\mathbf{x} \in \mathbb{I}^n,\ \sigma \in \Pi_{[n]}).$$

The symmetry property means that the aggregated value does not depend on the order of the arguments. This is required when combining criteria of equal importance or anonymous expert's opinions, e.g., symmetry is more natural in voting procedures than in multicriteria decision making, where criteria usually have different importances. Many aggregation functions introduced till now are symmetric. For example, AM, GM, OWA_ω are symmetric functions. Prominent examples of non-symmetric aggregation functions are weighted arithmetic means WAM_ω. In situations when judges, criteria, or individual opinions are not equally important, the symmetry property must be omitted.

Definition 3 $\mathsf{A} : \mathbb{I}^n \to \mathbb{I}$ *is an idempotent function if*

$$\mathsf{A}(n \cdot x) = x \qquad (x \in \mathbb{I}).$$

Idempotency is in some areas supposed to be a genuine property of aggregation functions, e.g., in multi-criteria decision making [5], where it is commonly accepted that if all criteria are satisfied at the same degree x then also the global score should be x. It is obvious that AM, WAM_ω, OWA_ω, Min, Max, and Med are idempotent functions, while Σ and Π are not. An element $x \in \mathbb{I}$ is an *idempotent* for $\mathsf{A} : \mathbb{I}^n \to \mathbb{I}$ if $\mathsf{A}(n \cdot x) = x$. In $[0,1]^n$ the product Π has no other idempotent elements than the extreme elements 0 and 1. As an example of an aggregation function in $[0,1]^n$ which is not idempotent but has a non-extreme idempotent element, take an arbitrarily chosen element $c \in\ ]0,1[$ and define the aggregation function $\mathsf{A}_{\{c\}} : [0,1]^n \to [0,1]$ as follows:

$$\mathsf{A}_{\{c\}}(x_1, \ldots, x_n) = \max\Big(0, \min\big(1, c + \sum_{i=1}^{n}(x_i - c)\big)\Big).$$

It is easy to see that the only idempotent elements for $\mathsf{A}_{\{c\}}$ are 0, 1, and c.

The next class of properties concern the "clustering" character of the aggregation functions, i.e., we assume that it is possible to partition the set of the arguments into disjoint subgroups, build the partial aggregation for each subgroup and then combine these partial results to get the global value. This condition may take several forms. The strongest one we will present is associativity. Other weaker formulations will also be presented, namely decomposability, autodistributivity, bisymmetry, self-identity.

Definition 4 $\mathsf{A} : \mathbb{I}^2 \to \mathbb{I}$ *is associative if, for all* $\mathbf{x} \in \mathbb{I}^3$, *we have*

$$\mathsf{A}\big(\mathsf{A}(x_1, x_2), x_3\big) = \mathsf{A}\big(x_1, \mathsf{A}(x_2, x_3)\big). \tag{1}$$

As examples of associative functions recall Min, Max, Σ, Π, P_F, P_L. Functions like AM and GM are not associative. In fact, associativity is a very strong and rather restrictive property, especially together with continuity. Therefore sometimes some modifications of associativity preserving its advantages (from the computational point of view) and extending the freedom in the choice of $\mathsf{A}^{(n)}$, $n > 2$, are introduced.

Definition 5 $\mathsf{A} : \cup_{n \geqslant 1} \mathbb{I}^n \to \mathbb{I}$ *is decomposable if* $\mathsf{A}(x) = x$ *for all* $x \in \mathbb{I}$ *and if*

$$\mathsf{A}(x_1, \ldots, x_k, x_{k+1}, \ldots, x_n) = \mathsf{A}\big(k \cdot \mathsf{A}(x_1, \ldots, x_k), (n-k) \cdot \mathsf{A}(x_{k+1}, \ldots, x_n)\big)$$

for all integers $0 \leqslant k \leqslant n$, *with* $n \geqslant 1$, *and all* $\mathbf{x} \in \mathbb{I}^n$.

By considering $k = 0$ (or $k = n$), we see that any decomposable function is range-idempotent. It follows that decomposability means that each element of any subset of consecutive elements from $\mathbf{x} \in \mathbb{I}^n$ can be replaced with their partial aggregation without changing the global aggregation. Decomposability also implies that the global aggregation does not change when altering some consecutive values without modifying their partial aggregation.

Definition 6 $\mathsf{A} : \mathbb{I}^2 \to \mathbb{I}$ *is bisymmetric if for all* $\mathbf{x} \in \mathbb{I}^4$, *we have*

$$\mathsf{A}\big(\mathsf{A}(x_1, x_2), \mathsf{A}(x_3, x_4)\big) = \mathsf{A}\big(\mathsf{A}(x_1, x_3), \mathsf{A}(x_2, x_4)\big).$$

In a certain respect, it has been used as a substitute for associativity and also for symmetry.

Depending on the kind of scale which is used, allowed operations on values are restricted. For example, aggregation on ordinal scales should be limited to operations involving comparisons only, such as medians and order statistics, while linear operations are allowed on interval scales. To be precise, a *scale of measurement* is a mapping that assigns real numbers to objects being measured. Stevens defined the *scale type* of a scale by giving a class of *admissible transformations*, transformations that lead from one acceptable scale to another.

The neutral element is again a well-known notion coming from the area of binary operations.

Definition 7 *Let* $\mathsf{A} : \cup_{n\geqslant 1}\mathbb{I}^n \to \mathbb{I}$ *be an aggregation function. An element* $e \in \mathbb{I}$ *is called a neutral element of* A *if, for any* $i \in \{1, \ldots, n\}$ *and any* $\mathbf{x} \in \mathbb{I}^n$ *such that* $x_i = e$*, then*

$$\mathsf{A}(x_1, \ldots, x_n) = \mathsf{A}(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n).$$

The neutral element can be omitted from aggregation inputs without influencing the final output. In multi-criteria decision making, assigning a score equal to the neutral element (if it exists) to some criterion means that only the other criteria fulfillments are decisive for the global evaluation.

Further properties of aggregation functions are related to additivity and its modifications.

Definition 8 $\mathsf{A} : \mathbb{I}^n \to \mathbb{I}$ *is*

(i) additive if

$$\mathsf{A}(\mathbf{x} + \mathbf{x}') = \mathsf{A}(\mathbf{x}) + \mathsf{A}(\mathbf{x}')$$

for all $\mathbf{x}, \mathbf{x}' \in \mathbb{I}^n$ *such that* $\mathbf{x} + \mathbf{x}' \in \mathbb{I}^n$*;*

(ii) is minitive if

$$\mathsf{A}(\mathbf{x} \wedge \mathbf{x}') = \mathsf{A}(\mathbf{x}) \wedge \mathsf{A}(\mathbf{x}')$$

for all $\mathbf{x}, \mathbf{x}' \in \mathbb{I}^n$*;*

(iii) is maxitive if

$$\mathsf{A}(\mathbf{x} \vee \mathbf{x}') = \mathsf{A}(\mathbf{x}) \vee \mathsf{A}(\mathbf{x}')$$

for all $\mathbf{x}, \mathbf{x}' \in \mathbb{I}^n$*.*

We now present the concept of *comonotonicity*. Two vectors $\mathbf{x}, \mathbf{x}' \in \mathbb{I}^n$ are said to be *comonotonic* if there exists a permutation $\pi \in \Pi_{[n]}$ such that

$$x_{\pi(1)} \leq \cdots \leq x_{\pi(n)} \quad \text{and} \quad x'_{\pi(1)} \leq \cdots \leq x'_{\pi(n)}.$$

Thus π orders the components of $\mathbf{x}$ and $\mathbf{x}'$ simultaneously. Another way to say that $\mathbf{x}$ and $\mathbf{x}'$ are comonotonic is that $(x_i - x_j)(x'_i - x'_j) \geq 0$ for every $i, j \in \{1, \ldots, n\}$. Thus if $x_i < x_j$ for some i, j then $x'_i \leq x'_j$.

Definition 9 $\mathsf{A} : \mathbb{I}^n \to \mathbb{R}$ *is comonotonic additive if*

$$\mathsf{A}(\mathbf{x} + \mathbf{x}') = \mathsf{A}(\mathbf{x}) + \mathsf{A}(\mathbf{x}')$$

for all comonotonic vectors $\mathbf{x}, \mathbf{x}' \in \mathbb{I}^n$ *such that* $\mathbf{x} + \mathbf{x}' \in \mathbb{I}^n$*.*

3 Classes of aggregation functions

Usually aggregation functions are divided into three classes, each possessing very distinct behavior: *conjunctive* functions, *disjunctive* functions and *internal* functions.

Definition 10 $\mathsf{A} : \mathbb{I}^n \to \mathbb{I}$ *is conjunctive if*

$$\mathsf{A}(\mathbf{x}) \leqslant \min x_i \qquad (\mathbf{x} \in \mathbb{I}^n).$$

Conjunctive functions combine values as if they were related by a logical "and" operator. That is, the result of combination can be high only if all the values are high. t-norms are suitable functions (defined on $[0,1]^n$) for doing conjunctive aggregation. However, they generally do not satisfy properties which are often requested for multicriteria aggregation, such as idempotence, scale invariance, etc.

Definition 11 $\mathsf{A} : \mathbb{I}^n \to \mathbb{I}$ *is disjunctive if*

$$\mathsf{A}(\mathbf{x}) \geqslant \max x_i \qquad (\mathbf{x} \in \mathbb{I}^n).$$

Disjunctive functions combine values as an "or" operator, so that the result of combination is high if at least one value is high. Such functions are, in this sense, dual of conjunctive functions. The most common disjunctive functions are t-conorms (defined on $[0,1]^n$). As t-norms, t-conorms do not possess suitable properties for criteria aggregation.

Definition 12 $\mathsf{A} : \mathbb{I}^n \to \mathbb{I}$ *is internal if*

$$\min x_i \leqslant \mathsf{A}(\mathbf{x}) \leqslant \max x_i \qquad (\mathbf{x} \in \mathbb{I}^n).$$

Between conjunctive and disjunctive functions, there is a third category, namely *internal* functions. They are located between min and max, which are the bounds of the t-norm and t-conorm families. In multicriteria decision aid, these functions are also called *compensative* operators. In fact, in this kind of functions, a bad (resp. good) score on one criterion can be compensated by a good (resp. bad) one on another criterion, so that the result of the aggregation will be medium.

4 Continuity

We consider now the usual *continuity* of aggregation functions and its strengthenings and weakenings.

Definition 13 $\mathsf{F} : \mathbb{I}^n \to \mathbb{R}$ *is a continuous function if*

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \mathsf{F}(\mathbf{x}) = \mathsf{F}(\mathbf{x}_0) \qquad (\mathbf{x}, \mathbf{x}_0 \in \mathbb{I}^n).$$

The continuity property means that the small change of arguments (possible minor error) will not cause a big output difference (output error). For non-decreasing functions continuity can be characterized alternatively, see [9].

Proposition 14 *For a non-decreasing function* $\mathsf{F} : \mathbb{I}^n \to \mathbb{R}$ *the following are equivalent*

(i) F *is continuous;*

(ii) F *is continuous in each coordinate, i.e., for every* $\mathbf{x} \in \mathbb{I}^n, i \in [n]$, *functions* $\mathsf{F}_{\mathbf{x},i} : I \to \mathbb{R}$ *given by*

$$\mathsf{F}_{\mathbf{x},i}(u) = \mathsf{F}(x_1, \ldots, x_{i-1}, u, x_i, \ldots, x_n)$$

are continuous;

(iii) F *has the intermediate value property, i.e., for all* $\mathbf{x}, \mathbf{y} \in \mathbb{I}^n$ *with* $\mathbf{x} \leqslant \mathbf{y}$ *and all* $c \in [\mathsf{F}(\mathbf{x}), \mathsf{F}(\mathbf{y})]$, *there exists* $\mathbf{z} \in \mathbb{I}^n$, *with* $\mathbf{x} \leqslant \mathbf{z} \leqslant \mathbf{y}$, *such that* $\mathsf{F}(\mathbf{z}) = c$ *(here* $\leqslant$ *denotes the classical Cartesian product partial order).*

Stronger forms of continuity
Uniformly continuous functions

Definition 15 *Let* $\| \cdot \| : \mathbb{R}^n \to \mathbb{R}^+$ *be a norm and* $D \subseteq \mathbb{R}^n$. *A function* $\mathsf{F} : D \to \mathbb{R}$ *is called uniformly continuous on* D *if for every* $\varepsilon > 0$ *there exists* $\delta > 0$ *such that* $|\mathsf{F}(\mathbf{x}) - \mathsf{F}(\mathbf{y})| < \varepsilon$, *whenever* $\|\mathbf{x} - \mathbf{y}\| < \delta$ *and* $\mathbf{x}, \mathbf{y} \in D$.

Absolutely continuous functions

Definition 16 *We say that* $\mathsf{F} : [a, b] \to \mathbb{R}$ *is absolutely continuous if for every* $\varepsilon > 0$ *there exists a* $\delta > 0$ *such that for any finite system of pairwise non-intersecting intervals* $]a_i, b_i[\subset]a, b[, i = 1, \ldots, n$, *for which* $\sum_{i=1}^{n}(b_i - a_i) < \delta$ *the inequality* $\sum_{i=1}^{n} |\mathsf{F}(b_i) - \mathsf{F}(a_i)| < \varepsilon$ *holds.*

Every absolutely continuous function on a closed interval is continuous on this interval. The opposite implication is not true, e.g., the function

$$\mathsf{F}(x) = \begin{cases} x \sin \frac{1}{x} & \text{if } x \in]0, 1] \\ 0 & x = 0, \end{cases}$$

is continuous on $[0, 1]$, but is not absolutely continuous on it.

The fundamental theorem of calculus is a part of the following

Theorem 17 *Let* $\mathsf{F} : [a, b] \to \mathbb{R}$. *Then the following are equivalent:*

(i) There is an integrable real-valued function f *such that*

$$\mathsf{F}(x) = \mathsf{F}(a) + \int_a^x f(t)\, dt$$

for every $x \in [a, b]$.

(ii) $\int_a^x \mathsf{F}'(t)\,dt$ *exists and is equal to* $\mathsf{F}(x) - \mathsf{F}(a)$ *for every* $x \in [a, b]$.

(iii) F *is absolutely continuous.*

We extend the definition of absolute continuity on functions of more variables using Theorem 17.

Definition 18 *A function* $\mathsf{F} : \mathbb{I}^n \to \mathbb{R}$ *is absolutely continuous if and only if its partial derivative* $\frac{\partial \mathsf{F}(x_1,\ldots,x_n)}{\partial x_1 \ldots \partial x_n}$ *exist almost everywhere and for any fixed* $(a_1, \ldots, a_n)$ *from the interior of* $\mathbb{I}^n$,

$$\mathsf{F}(x_1, \ldots, x_n) = \int_{(a_1,\ldots,a_n)}^{(x_1,\ldots,x_n)} f(x_1, \ldots, x_n)\, dx_1 \ldots dx_n + \mathsf{F}(a_1, \ldots, a_n),$$

where

$$f(x_1, \ldots, x_n) = \frac{\partial \mathsf{F}(x_1, \ldots, x_n)}{\partial x_1 \ldots \partial x_n}$$

if the partial derivative exists, and else $f(x_1, \ldots, x_n) = 0$.

Lipschitz condition

The continuity property can be strengthened into the well-known Lipschitz condition.

Definition 19 *Let* $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}^+$ *be a norm. If a function* $\mathsf{F} : \mathbb{I}^n \to \mathbb{R}$ *satisfies the inequality*

$$|\mathsf{F}(\mathbf{x}) - \mathsf{F}(\mathbf{y})| \leqslant c\|\mathbf{x} - \mathbf{y}\| \quad (\mathbf{x}, \mathbf{y} \in \mathbb{I}^n), \tag{2}$$

$c \in]0, \infty[$, *then we say that* F *satisfies Lipschitz condition (with respect to* $\|\cdot\|$*). The greatest lower bound of constants* $c > 0$ *in (2) is the Lipschitz constant.*

Remark 20 *If a function* $\mathsf{F} : \mathbb{I}^n \to \mathbb{R}$ *satisfies the inequality*

$$|\mathsf{F}(\mathbf{x}) - \mathsf{F}(\mathbf{y})| \leqslant c\|\mathbf{x} - \mathbf{y}\|^\alpha \quad (\mathbf{x}, \mathbf{y} \in \mathbb{I}^n),$$

where $0 < \alpha \leq 1$ *and* $c \in]0, \infty[$, *then we say that* F *satisfies Lipschitz condition (with respect to* $\|\cdot\|$*) of order* α. *In the case* $0 < \alpha < 1$ *the condition (2) is also called Hölder condition of order* α.

As an important example we have the norm $\|\mathbf{x}\|_p := \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$, for some $p \in [1, \infty[$, and $\|\mathbf{x}\|_\infty := \max |x_i|$ is the Chebyschev norm. By convention, when the norm $\|\cdot\|$ is not specified, the L^1 norm on $\mathbb{R}^n$, i.e., $\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|$, is taken into account. Moreover, we can stress the actual value of constant c in (2) when speaking that it is *c-Lipschitz*. c-Lipschitzianity of aggregation function allows to estimate the relative output error in comparison with input errors

$$|\mathsf{F}(\mathbf{x}) - \mathsf{F}(\mathbf{y})| \leq c\varepsilon$$

whenever $\|\mathbf{x}-\mathbf{y}\| \le \varepsilon$ for some $\varepsilon > 0$. As a consequence of the fact that we have for $p \in [1,\infty[$ the inequalities

$$\|\mathbf{x}\|_p \le \|\mathbf{x}\|_1 \le n^{1-\frac{1}{p}}\|\mathbf{x}\|_p, \quad \text{and } \|\mathbf{x}\|_1 \le n\|\mathbf{x}\|_\infty \le n\|\mathbf{x}\|_p$$

and

$$\|\mathbf{x}\|_q \le \|\mathbf{x}\|_p \text{ whenever } q \ge p$$

for all $\mathbf{x} \in \mathbb{I}^n$ we obtain the following result.

The Lipschitz property of functions is defined standardly on domains where the norm cannot achieve the value infinity. Formally it can be defined also on $\mathbb{I}^n$ for an unbounded interval $\mathbb{I}$, however, if $\mathbb{I}$ is closed then the Lipschitz property does not imply continuity, in general. For example, the weakest aggregation function F on $[0,\infty]^n$ is Lipschitz for any norm, however, it is not continuous.

In all cases the Lipschitz property implies the absolute continuity, so it is stronger property, e.g., $\log x$ on $]0,1]$ is absolutely continuous but not Lipschitz of any order. Also, we have the following connection.

Proposition 21 *If, in Definition 16 of an absolutely continuous function, the requirement that the pairwise intersections of intervals are empty is discarded, then the function will satisfy Lipschitz condition with some constant.*

Proposition 22 *Let $[a,b]$ be a bounded interval. The smallest and the greatest aggregation function 1-Lipschitz with respect to the norm $\|\cdot\|_p$ are given by $\mathsf{A}_*^{(n)} : [a,b]^n \to [a,b]$,*

$$\mathsf{A}_*^{(n)}(\mathbf{x}) := \max(b - \|nb - \mathbf{x}\|_p, a),$$

and $\mathsf{A}^{(n)*} : [a,b]^n \to [a,b]$,

$$\mathsf{A}^{(n)*}(\mathbf{x}) := \min(a + \|\mathbf{x} - na\|_p, b),$$

respectively.

Each Lipschitz function (with respect to any norm $\|\cdot\|$) for $\mathbb{I}$ not closed infinite interval is continuous. The converse is false in general.

Example 23 *The geometric mean GM on the interval $[0,1]^n$ or on the interval $[0,\infty[^n$ is a continuous function which is not Lipschitz.*

Example 24 *The arithmetic mean $\mathsf{AM} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{R}$ is 1-Lipschitz independently of the interval $\mathbb{I}$. Extended aggregation function $\mathsf{Q} : \cup_{n\in\mathbb{N}}[0,1]^n \to [0,1]$ given by*

$$\mathsf{Q}(x_1,\ldots,x_n) := \Pi_{i=1}^n x_i^i$$

is not Lipschitz, though each $\mathsf{Q}^{(n)}$ is Lipschitz (the best Lipschitz constant c_n for $\mathsf{Q}^{(n)}$ is $c_n = n$).

Weaker forms of continuity

Definition 25 $\mathsf{F} : \mathbb{I}^n \to \mathbb{R}$ *is called lower semi-continuous or left-continuous if, for all* $(\mathbf{x}^{(k)})_{k\in\mathbb{N}} \subset (\mathbb{I}^n)^{\mathbb{N}}$ *such that* $\vee_k \mathbf{x}^{(k)} \in \mathbb{I}^n$*, it holds*

$$\bigvee_k \mathsf{F}(\mathbf{x}^{(k)}) = \mathsf{F}\left(\bigvee_k \mathbf{x}^{(k)}\right).$$

Definition 26 $\mathsf{F} : \mathbb{I}^n \to \mathbb{R}$ *is called upper semi-continuous or right continuous if, for all* $(\mathbf{x}^{(k)})_{k\in\mathbb{N}} \subset (\mathbb{I}^n)^{\mathbb{N}}$ *such that* $\wedge_k \mathbf{x}^{(k)} \in \mathbb{I}^n$*, it holds*

$$\bigwedge_k F(\mathbf{x}^{(k)}) = F\left(\bigwedge_k \mathbf{x}^{(k)}\right).$$

5 Non-additive measures

Let us consider $\mathbb{I} = [0,1]$ and $N = \{1,\dots,n\}$. A *set function* m on N is a function from 2^N to $\mathbb{R}$. We have already noted above the bijection between vertices of $[0,1]^n$ and subsets of N. Hence a subset $A \subseteq N$ is equivalently denoted by $(1_A, 0_{A^c}) \in [0,1]^n$, or by its characteristic function $\mathbf{1}_A$ defined over N. We denote $\mathbf{x} = (x_1,\dots,x_n)$. Using the above equivalence, any set function m bijectively corresponds to a *pseudo-Boolean function* $f_m : \{0,1\}^n \to \mathbb{R}$ by $f_m(x) = m(A_x)$ for all $\mathbf{x} \in \{0,1\}^n$, where $A_{\mathbf{x}} = \{i \in N \mid x_i = 1\}$. Conversely, to any pseudo-Boolean function f corresponds a unique set function m_f such that $m_f(A) := f(1_A, 0_{A^c})$. Pseudo-Boolean functions are widely used in operations research. Cooperative game theory is devoted to a particular class of set functions, called *transferable utility games in characteristic form*. We will call them *games* or *non-additive measure* for simplicity. In the context of game theory, the set N is the set of players. A game $m : 2^N \to \mathbb{R}$ is a set function satisfying $m(\varnothing) = 0$. Useful examples of games are *unanimity games*. For any $A \subseteq N$, the *unanimity game* u_A on N is defined by:

$$u_A(B) := \begin{cases} 1, & \text{if } B \supseteq A \\ 0, & \text{otherwise.} \end{cases}$$

Note that $u_\varnothing$ is not a game since $u_\varnothing(\varnothing) = 1$.

A *capacity* $m : 2^N \to \mathbb{R}_+$ is a game such that $\mu(A) \leq \mu(B)$ whenever $A \subseteq B$ (*monotonicity*). A capacity is *normalized* if $\mu(N) = 1$. Capacities are monotonic games, and were introduced originally by Choquet in 1953. They were rediscovered by Sugeno in 1974 under the name *fuzzy measure*.

Important connection with aggregation functions can be described in the following way. Suppose we use as an aggregation function the weighted arithmetic mean

$$\mathsf{WAM}_{\mathbf{w}}(\mathbf{x}) = \frac{w_1 x_1 + \cdots + w_n x_n}{w_1 + \cdots + w_n}$$

with respect to some weight vector $\mathbf{w} \in [0,1]^n$. It is easy to relate $\mathbf{w}$ to the values taken on by $\mathsf{WAM}_{\mathbf{w}}$, using particular vectors in $[0,1]^n$, namely $\mathbf{1}_i$: $\mathsf{WAM}_{\mathbf{w}}(\mathbf{1}_i) = w_i$ for all $i \in N$. This means that the value of function $\mathsf{WAM}_{\mathbf{w}}$ on $[0,1]^n$ is solely determined by its value at the endpoints of the n dimensions, which represents the weight of each dimension. In fact, the exact way $\mathsf{WAM}_w(x)$, $x \in [0,1]$, is determined from $\mathsf{WAM}_w(\mathbf{1}_i)$, $i = 1, \ldots, n$, is linear interpolation. One may construct more complicated aggregation functions A by using more points in $[0,1]^n$ to determine A. A natural yet simple choice would be to take all vertices of $[0,1]^n$, namely $\{\mathbf{1}_A\}_{A \subseteq N}$. These include the previous endpoints of dimensions. Doing so, we have defined a set of weights $\{w_A\}_{A \subseteq N}$, by

$$\mathsf{A}_w(\mathbf{1}_A) = w_A, \quad A \subseteq N.$$

It remains to construct A_w on $[0,1]^n$ by some means (e.g., linear interpolation), using these points. By analogy with the previous case, w_A is the weight of the subset A of dimensions.

In the case of WAM_w, the weight vector had no peculiar property, beside non-negativeness and normalization $\sum_i w_i = 1$. If weights are assigned to subsets of dimensions, then some properties are natural, especially if dimensions represent criteria or attributes, or individuals (voters, experts). In this framework, $x \in [0,1]^n$ is a vector of scores, and $\mathsf{A}_w(x)$ is the aggregated overall score, reflecting the score of each criterion or individual. Hence, $\mathsf{A}_w(1_A, 0_{A^c})$ is the overall score of an object having the maximal score for all criteria (individuals) in A and the minimal score otherwise, so that the following properties are natural:

(i) $w_\varnothing = 0$, since the object $(1_\varnothing, 0_N)$ is the worst possible;

(ii) $w_N = 1$, since the object $(1_N, 0_\varnothing)$ is the best possible;

(iii) $w_A \leq w_B$ whenever $A \subseteq B$, since object $(1_B, 0_{B^c})$ is at least better on one dimension than $(1_A, 0_{A^c})$.

Considering w as a set function on N, what we have defined above is nothing else than a *capacity*.

Let m be a set function on N, i.e., an element of $\mathbb{R}^{2^N}$. A *transform* is any mapping $T : \mathbb{R}^{2^N} \to \mathbb{R}^{2^N}$. The transform is *linear* if for any m_1, m_2 in $\mathbb{R}^{2^N}$ and any $\lambda_1, \lambda_2 \in \mathbb{R}$ it holds $T(\lambda_1 m_1 + \lambda_2 m_2) = \lambda_1 T(m_1) + \lambda_2 T(m_2)$, and it is *invertible* if T^{-1} exists. There are several useful invertible linear transforms of set functions. The best known one is the *Möbius transform*

$$\mu(A) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} m(B).$$

μ is said to be the *Möbius transform* (or *Möbius inverse*) of m. It is a linear and invertible transform, and $\mu(\varnothing) = m(\varnothing)$. The Möbius transform has been rediscovered many times. In the field of pseudo-Boolean functions, it appears as coefficients in the multilinear polynomial form of any pseudo-Boolean function f:

$$f(x) = \sum_{T \subseteq N} \left[a_T \prod_{i \in T} x_i \right], \quad \forall x \in \{0,1\}^n.$$

In the field of cooperative game theory was found by Shapley in the form

$$\mu(A) = \sum_{B \subseteq N} m_B u_B(A), \quad \forall A \subseteq N,$$

i.e., any game (in fact, any set function) can be expressed in a unique way by unanimity games.

6 The Benvenuti integral

We consider a general integral with respect to a capacity, which cover many well-known integrals. Benvenuti integral is based on the chain representation (comonotone representation) of input vectors and two binary operations $\oplus$ and $\odot$. For a constant $b \in \,]0,\infty]$, operation $\oplus : [0,b]^2 \to [0,b]$ is supposed to be a continuous t-conorm, i.e., an associative continuous binary aggregation function with neutral element 0. For another constant $c \in \,]0,\infty]$ (the case $c = b$ is possible and most frequent case) operation $\odot : [0,b] \times [0,c] \to [0,b]$ is a non-decreasing binary operation which is right-distributive with respect to $\oplus$, i.e.,

$$(u \oplus v) \odot w = (u \odot v) \oplus (v \odot w)$$

for all $u, v \in [0,b]$ and $w \in [0,c]$. Moreover, define a binary operation $\ominus : [0,b]^2 \to [0,b]$ associated to $\oplus$ by

$$u \ominus v = \inf\{t \in [0,b] \mid v \oplus t \geq u\}.$$

Definition 27 *For a fixed $n \in \mathbb{N}$, let $m : 2^N \to [0,c]$ be a monotone set function (capacity).* Benvenuti integral $B_m^{\oplus,\odot} : [0,b]^n \to [0,b]$ *is given by*

$$B_m^{\oplus,\odot}(\mathbf{x}) := \oplus_{i=1}^n (x_{(i)} \ominus x_{(i-1)}) \odot m(A_{(i)}),$$

with σ a permutation on N such that $x_{\sigma(1)} \leq x_{\sigma(2)} \leq \cdots \leq x_{\sigma(n)}$, with the convention $x_{\sigma(0)} := 0$, and $A_{\sigma(i)} := \{\sigma(i), \ldots, \sigma(n)\}$.

Many important special cases are in the following example.

Example 28 *(i) Let $b = c = \infty, \oplus = +, \odot = \cdot$ on $[0,\infty]$. Then*

$$B_m^{+,\cdot}(\mathbf{x}) = \mathcal{C}_m(\mathbf{x}) = \sum_{i=1}^n (x_{\sigma(i)} - x_{\sigma(i-1)}) m(A_{\sigma(i)}),$$

reduces on the Choquet integral.

(ii) *Let* $b = c = 1, \oplus = \vee, \odot = \wedge$. *Then*

$$B_m^{\vee,\wedge}(\mathbf{x}) = \mathcal{S}_m(\mathbf{x}) := \bigvee_{i=1}^{n} \left(x_{\sigma(i)} \wedge m(A_{\sigma(i)})\right),$$

reduces on Sugeno integral.

(iii) *Let* $b = c = 1, \oplus = \mathsf{S}_\mathbf{P}$ *(probabilistic sum), i.e.,* $u \oplus v = u + v - uv$, *and* $\odot : [0,1]^2 \to [0,1]$ *is a uninorm generated by a multiplicative generator* $\varphi : [0,1] \to [0,1]$ *given by* $\varphi(x) = -\log(1-x)$, *i.e.,*

$$u \odot v = \exp(-\log(1-u)\log(1-v)),$$

and the neutral element of $\odot$ *is* $e = 1 - \exp(-1)$. *Then*

$$B_m^{\oplus,\odot}(\mathbf{x}) = \varphi^{-1}\left(C_{\varphi \circ m}(\varphi \circ \mathbf{x})\right),$$

i.e., Benvenuti integral is a φ*-transform of the Choquet integral.*

(iv) *For* $\oplus$*-measure* m *the integral* $B_m^{\oplus,\odot}$ *reduces on the pseudo-integral, see [11].*

(v) *Let* $b = c = 1, \oplus = \vee, \odot = \cdot$. *Then* $B_m^{\vee,\cdot}(\mathbf{x}) = \max_i(w_i \cdot x_i)$, *where* $x_i = f(i)$, *gives the Shilkret integral (where the integral was considered with respect to* $\mathsf{S}_\mathbf{M}$*-measures).*

Recall some general properties of Benvenuti integral.

Theorem 29 *(i) The Benvenuti integral* $B_m^{\oplus,\odot}$ *is monotone.*

(ii) *If* $b = c$ *and* $\odot$ *is associative, then* $B_m^{\oplus,\odot}$ *is* $\odot$*-homogeneous, i.e.,*

$$B_m^{\oplus,\odot}(a \odot u) = a \odot B_m^{\oplus,\odot}(u)$$

for all $a \in [0,b]$.

(iii) *If* $\odot$ *has a left neutral element, i. e.,* $e \odot u = u$ *for all* $u \in [0,b]$, *then*

$$B_m^{\oplus,\odot}(e \cdot \mathbf{1}_A) = m(A).$$

(iv) *If* $\odot$ *has a right neutral element, i. e.,* $u \odot e = u$ *for all* $u \in [0,b]$, *and* $m(N) = e$, *then*

$$B_m^{\oplus,\odot}(u) = u$$

for all $u \in [0,b]$.

For more details, especially concerning the Benvenuti integral on abstract spaces, we recommend [2].

Acknowledgement The work has been supported by the project MNZŽSS-144012 and the project "Mathematical Models for Decision Making under Uncertain Conditions and Their Applications" supported by Vojvodina Provincial Secretariat for Science and Technological Development.

References

[1] R. J. Aumann, L. S. Shapley: *Values of Non-Atomic Games* Princton Univ. Press, 1974.

[2] P. Benvenuti, R. Mesiar, D. Vivona: *Monotone Set Functions-Based Integrals.* In: E. Pap ed. Handbook of Measure Theory, Ch **33.**, Elsevier, 2002, 1329-1379.

[3] T. Calvo, G. Mayor, R. Mesiar (Eds.), *Aggregation Operators*, volume 40 of *Studies in Fuzziness*, Physica-Verlag, Heidelberg, 2002.

[4] D. Dubois, E. Pap, H. Prade, Hybrid probabilistic-possibilistic mixtures and utility functions, (Eds. B. De Baets, J. Fodor, P. Perny) Preferences and Decisions under Incomplete Knowledge , Springer-Verlag, 2000, 51-73.

[5] J. C. Fodor, M. Roubens, *Fuzzy Preference Modeling and Multi–Criteria Decision Aid*, Kluwer Academic Publisher, Dordrecht, 1994.

[6] M. Grabisch, T. Murofushi, M. Sugeno (Eds.), *Fuzzy Measures and Integrals-Theory and Applications*, volume 40 of *Studies in Fuzziness*, Physica-Verlag, Heidelberg, 2000.

[7] M. Grabisch, H. T. Nguyen, E. A. Walker, *Fundamentals of Uncertainity Calculi with Application to Fuzzy Inference*, Kluwer Academic Publishers, Dordrecht/Boston/London, 1995.

[8] M. Grabisch, J. L. Marichal, R. Mesiar, E. Pap, Aggregation Functions for Decision Making (book under preparation).

[9] E. P. Klement, R. Mesiar, E. Pap, *Triangular Norms*, Kluwer Academic Publishers, Dordrecht, 2000.

[10] R. Mesiar, Choquet–like integrals. *J. Math. Anal. Appl.* **194** (1995) , 477-488.

[11] E. Pap, *Null–Additive Set Functions*, Kluwer Academic Publishers, Dordrecht/Boston/London, 1995.

[12] E. Pap, Pseudo-analysis as a mathematical base for soft computing, *Soft Computing* **1** (1997), 61-68.

[13] E. Pap, Aggregation operators in the engineering design, (Eds. C. Tomasa, G. Mayor, R. Mesiar) volume 97 of *Studies in Fuzziness and Soft Computing*, Springer-Verlag, 2002, 195-223.

[14] Z. Wang, G. J. Klir, *Fuzzy Measure Theory*, Plenum Press, New York and London, 1992.